Observational, One-Arm Studies and Randomized Population-Based Trials for Evaluation of the Efficacy of Lung Cancer Screening

Eugenio Paci, MD

(*J Thorac Oncol.* 2007;2: Suppl 1, 45–46)

In 1999, the Early Lung Cancer Action Project (ELCAP) published the results of its evaluation of early detection of lung cancer in high-risk subjects in New York, showing that low-dose computed tomography (CT) is a sensitive tool for the identification of non-calcified pulmonary nodules and early lung cancer.[1] A methodological debate then ensued over the need to demonstrate by means of a randomized controlled trial (RCT) the reduction of lung cancer mortality for people included in CT scan screening versus those unscreened or screened by chest radiograph. In their recent article, the International ELCAP group (I-ELCAP) has confirmed (with a greater number of cases) that CT scan is a sensitive screening test for early lung cancer and has shown high survival rates among patients with stage I screen-detected disease.[2] Several other one-arm studies evaluating the screening test at prevalence and incidence screening have confirmed its good performance.[3,4]

CHEST RADIOGRAPH SCREENING RANDOMIZED TRIALS

Lung cancer screening by chest radiograph, with or without cytology testing, was introduced more than 30 years ago, and several RCTs reported the absence of mortality reduction at the end of follow-up. The conclusion of the most important international agencies is that chest radiograph is not suitable as a lung cancer screening test.[5] The most important RCT was the Mayo Lung Screening Project, and after 16 years of follow-up, an update reinforced its conclusions that mortality had not diminished.[6] More recently, the study has also been updated for incidence rates, and earlier findings have again been confirmed.[7] In particular, an excess of lung cancer cases detected in the intervention group suggests over-diagnosis. In the accompanying editorial to the Mayo Project's updated mortality results, Black[8] discussed in depth the issue of over-diagnosis bias in chest radiograph trials as the greater harm associated with lung cancer screening but was open to the possibility that CT scan screening could reduce mortality and be highly effective.

In several articles, Miettinen et al. harshly rebutted the conclusion about over-diagnosis and accused exponents of pseudoevidence and specious reasoning.[9]

THE METHODOLOGICAL DEBATE

After the critique by the ELCAP group of so-called "orthodoxy,"[10] a scientific discussion arose about the need for experimental evaluation. In RCTs, the endpoint is cause-specific mortality and the comparison between the mortality rates of the intervention and control arms. The assumption was that screening would not change the timing of death if early diagnosis and treatment did not change the natural history of the disease.

However, in a new perspective based on the experience of lung cancer screening trials, the RCT is no longer a black box design. Today, it is designed with detailed data for all the performance parameters, case characteristics, and (in the new generation trials) related studies of biomolecular markers, and possibly quality-of-life and costs.

Analysis of the large dataset of 31,567 screened subjects in the recent I-ELCAP publication showed very high survival rates for subjects with screen-detected tumors, low rates of interval cancers, and very poor diagnosis of unresected stage I lung cancer cases.[2]

There are several factors to be considered in the evaluation of survival rates as indicators of screening efficacy. The first of these factors is lead time. It explains survival benefit as a consequence of diagnostic anticipation in asymptomatic subjects. In the I-ELCAP study, survival for all tumor types and sizes detected at screening and in the interval between screenings was 80%; extremely high when compared with that observed in population-based series (10–15% at 5 years). Of the 484 patients with lung cancer, 412 were clinical stage I, and their estimated survival rate was comparable with that observed in stage I cancers at 8 years in the SEER cancer registries. However, although estimated over 10 years, the rates were based on few cases with long follow-up: the follow-up range was 1 to 123 months, and the median was 40 months.

Unit of Clinical and Descriptive Epidemiology, CSPO Scientific Institute for Oncological Prevention, Florence, Italy.
Disclosure: The author declares no conflict of interest.
Address for correspondence: Eugenio Paci, Unit of Clinical and Descriptive Epidemiology, CSPO, Via San Salvi, 12 50135 Florence, Italy. E-mail: e.paci@cspo.it

ISSN: 1556-0864/07/0205-0045

The second factor is the possibility of over-diagnosis, the detection at screening of cancer cases with long mean sojourn time. An excess of cases, with low probability of progression, would inflate survival rates for years. There is no generally accepted way to distinguish individual cancers with low and high probabilities of progression. There are two components of over-diagnosis in screening: the detection of slowly growing cancers and the detection of cancers in subjects who would have had low probability of dying of lung cancer in their lifetime because of competing causes of death. The third major factor to be discussed is prognostic selection. There is a greater probability of detecting good prognosis cancer in studies in which healthy volunteers are selected for screening, such as the I-ELCAP study, because fast-growing and more aggressive cancers are less likely to be represented.

THE CT SCAN RANDOMIZED SCREENING TRIALS

The Lung Screening Study was a randomized feasibility study designed as a pilot for a large RCT.[11] A total of 1660 subjects were randomized to CT scan prevalence screening and one repeat screening and 1658 subjects to chest radiograph. With a compliance rate of 96% at baseline and 86% at 1 year, 40 lung cancer cases were diagnosed in the intervention arm and 20 in the chest radiograph arm. Of the cases in the intervention arm, 48% were stage I, compared with 40% in the chest radiograph arm. This interim result is the only currently published result of a randomized screening trial.

Larger RCTs are underway in the United States and Europe. The National Cancer Institute (NCI) is conducting a large randomized trial enrolling more than 50,000 subjects (NSLT),[12] and the Nelson study[13] has enrolled approximately 18,000 subjects in the Netherlands and Belgium and 4000 in Denmark. The Italung-CT RCT,[14] which is ongoing in the Italian cities of Florence, Pisa, and Pistoia, has enrolled patients from general practitioners' lists: 1613 in the intervention and 1593 in the usual care arms. Subjects were men and women aged 55–69, smokers (all 20-pack years or more) or former smokers (who stopped smoking within 10 years) who were randomized to the active (CT scan for 4 years) or passive (usual care) arm after informed consent. The third round of screening is in progress.

All these studies have comparable study designs and aims—a collaborative effort in evaluating mortality reduction achieved by lung cancer screening. In a meeting held in Liverpool recently, several research groups of one-arm and randomized screening trials have agreed to continue their collaboration sharing data and clinical experience.

CONCLUSION

The aim of randomized screening trials comparing the experience of a group of subjects offered a screening regimen with a comparable, non-screened, or differently screened group is to demonstrate the reduction of mortality that would be achieved by the early diagnosis of lung cancers. The aim of one-arm studies is to evaluate screening performance in terms of cancer detection, interval cancer cases, tumor characteristics, and survival rates. The crucial difference between these experimental study types, as I argued in a methodological article,[15] lies not only in their aims but also in the availability of the control group experience, against whom the active arm's probability of screen detection, tumor characteristics, and lung cancer survival rates should be compared. Assuming that, in the absence of intervention, lung cancer will progress and kill the subject did not take into consideration the possibility of over-diagnosis—the screen detection of indolent, slow growing, possibly not growing, lung cancers—which was documented in the old screening trials.

Although there are commonalities in comparing one-arm and RCTs, there are also differences between studies. The differences lie in the pattern of risk of screened subjects, in the technology used, in recall and detection rates, and in positive predictive values (which may not be related to the type but to the specificity of each study design). Only by performing early comparative work among studies and by sharing the results of quality evaluations of screening performance will public health agencies, scientists, and advocates be reassured that the future results of RCTs will be comparable with the most important and successful one-arm screening trials. And, hopefully, the scientific community will avoid a war over screening such as has been seen in the field of breast cancer screening evaluation.

REFERENCES

1. Henschke CI, McCauley DI, Yankelevitz DF, et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. *Lancet* 1999;354:99–105.
2. International Early Lung Cancer Action Program Investigators, Henschke CI, Yankelevitz DF, Libby DM, Pasmantier MW, Smith JP, Miettinen OS. Survival of patients with stage I lung cancer detected on CT screening. *N Engl J Med* 2006;355:1763–1771.
3. Swensen SJ, Jett JR, Hartman TE, et al. CT screening for lung cancer: five-year prospective experience. *Radiology* 2005;235:259–265.
4. Pastorino U, Bellomi M, Landoni C, et al. Early lung-cancer detection with spiral CT and positron emission tomography in heavy smokers: 2-year results. *Lancet* 2003;362:593–597.
5. US Preventive Services Task Force. http://www.ahrq.gov/clinic/uspstfix.htm. Accessed January 1, 2007.
6. Marcus PM, Bergstralh EJ, Fagerstrom RM, et al. Lung cancer mortality in the Mayo Lung Project: impact of extended follow-up. *J Natl Cancer Inst* 2000;92:1308–1316.
7. Marcus PM, Bergstralh EJ, Zweig MH, Harris A, Offord KP, Fontana RS. Extended lung cancer incidence follow-up in the Mayo Lung Project and overdiagnosis. *J Natl Cancer Inst* 2006;98:748–756.
8. Black WC. Overdiagnosis: an underrecognized cause of confusion and harm in cancer screening. *J Natl Cancer Inst* 2000;92:1280–1282.
9. Miettinen OS, Henschke CI. CT screening for lung cancer: coping with nihilistic recommendations. *Radiology* 2001;221:592–596.
10. Miettinen OS. Screening for lung cancer: do we need randomized trials? *Cancer* 2000;89(11 Suppl):2449–2452.
11. Gohagan JK, Marcus PM, Fagerstrom RM, et al. The Lung Screening Study Research Group. Final results of the Lung Screening Study, a randomized feasibility study of spiral CT versus chest X-ray screening for lung cancer. *Lung Cancer* 2005;47:9–15.
12. NCI Lung Screening Trial. http://www.cancer.gov/nlst. Accessed January 1, 2007.
13. van Iersel CA de Koning HJ, Draisma G, et al. Risk-based selection from the general population in a screening trial: Selection criteria, recruitment and power for the Dutch-Belgian randomised lung cancer multi-slice CT screening trial (NELSON). *Int J Cancer* 2007;120:868–874.
14. Picozzi G, Paci E, Lopez Pegna A, et al. Screening of lung cancer with low dose spiral CT: results of a three year pilot study and design of the randomised controlled trial "Italung-CT". *Radiol Med (Torino)* 2005; 109:17–26.
15. Paci E. Lung cancer screening: the methodological debate. *Lung Cancer* 2002;38:S17–21.